



SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Bhat, Ramesh, A.
Henderson,, Ruth
Hsiao, Chulai
Karathanasis, Sotirios
- (ii) TITLE OF THE INVENTION: NOVEL HUMAN ESTROGEN RECEPTOR-BETA
- (iii) NUMBER OF SEQUENCES: 13
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESS: Darby & Darby, P.C.
(B) STREET: 805 Third Avenue, 27th Floor
(C) CITY: New York
(D) STATE: NY
(E) COUNTRY: USA
(F) ZIP: 10022
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 09/429,832
(B) FILING DATE: 29-OCT-1999
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/906365
(B) FILING DATE: 03- AUG-1997
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Robinson, Joseph R
(B) REGISTRATION NUMBER: 33,446
(C) REFERENCE/DOCKET NUMBER: 0646/1D205
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 212-527-7700
(B) TELEFAX: 212-753-6237
(C) TELEX: 236687

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1086 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CAGCCATTAT	ACTTGCCTCAG	GAATCTTTTGA	GAACATTATA	ATGACCTTTG	TGCTCTTTCT	60
PGJAAGGTGT	TTTCTCAGCT	GCTATCTCAA	SACATGSATA	TAAAAAATC	ACCATCTAGC	120
CTTAATTCCT	CTTCCCTCTA	CAACTTCAAT	CAATCCATCT	TACCCCTTGA	GCACGCTCTC	180
ATATACATAA	CTTCCCTCTA	TGTAGACAGC	CACCATGAAT	ATCCAGCCAT	SACATTTCTAT	240
AGCCCTGCTG	TGATGAATTA	CAGCATTCCT	AGCAATGTCA	CTAACTTGA	AGGTGGGCT	300
GGTGGGCAGA	CCACAAGCTC	AAATGTGTTT	TGGCCAACAC	CTGGGACCT	TTCTCTTTTA	360
GTGGTCCATC	GCCAGTTATC	ACATCTGTAT	GGGAACCTC	AAAAAGTCC	CTGGTGTGAA	420
GCAGATCTC	TAGAACACAC	CTTACCTGTA	AACAGAGAGA	CACTGAAAA	GAAGTTTAT	480
GGGAACCTTT	GCCGCAAGCT	TGTTACTGTT	CCAGGTTCAA	AGAGGGATGC	TCACTTCTGC	540
GCTGTCTGCA	GCGATTACGC	ATCAGGATAT	CACATATGGAG	TCTGCTCTGT	TGAAGGATGT	600
AAAGCCCTTT	TTAAAAAGAG	CATTCAAGGA	CATAATGATT	ATATTTGTCC	AGCTACAAAT	660
CAGTGTACAA	TGATAAAAA	CGGCGCAAG	AGCTGCCAG	CTGTGCACT	TGGGAAGTGT	720
TACGAAGTGG	GAATGTGAA	GTGTGGCTCC	CGGAGAGAGA	GATGTGGTA	CTGGCTTGGG	780
CGGAGACAGA	GAAGTGGCGA	CGAGCAGCTC	CAGTGTGCGG	GCAAGGGCAA	GAGGAAGTGC	840
GGCCACGCGC	CCGAGTGGCG	GGAGCTGCTC	CTGGAGGCCC	TGAGCCCGCA	GCAGCTAGTG	900
CTCACCTTCC	TGGAGGCTGA	GGGCTCCAT	GTGCTGATCA	GGCGCCCGAG	TGGGCTCTTC	960
ACCGAGGCTT	CCATGATGAT	GTCTCTGACC	AAGTTGGCCG	ACAAAGAGTT	GCTACACATG	1020
ATCAGCTGGG	CCAAAGAGAT	TCTGGCTTTT	CTGGAGCTCA	GCTGTGTTGA	CCAGGTGGCG	1080
CTCTTGGAAG	GCTGTGGGAT	GGAGGTGTTA	ATGATGGGGC	TGATGTGGCG	CTCAATTGAC	1140
CACCCCGGCA	AGCTCATCTT	TGCTCCAGAT	CTTGTTCTTG	ACAGGGATGA	GGGGAAATGC	1200
GTAGAAGGAA	TTCTGAAAT	CTTTGACATG	CTCTGGGCAA	CTACTTCAAG	GTTTCGAGAG	1260
TTAAAACTCC	AACACAAAGA	ATATCTCTGT	GTCAAGGCCA	TGATCTGCT	CAATTCCAST	1320
ATGTACCTTC	TGGTCAAGC	GACCCAGGAT	GCTGACAGCA	GCCGGAAGCT	GGCTCACTTG	1380
CTGAACGCGC	TGACCGATGC	TTTGCTTTTG	GTGATTGCCA	AGAGGGGCAT	CTCTGCCAG	1440
CAGCAATCCA	TGCGCTTGGC	TAACCTCTCT	ATGCTCTCT	CCGACCTCAG	GCATGCGAGT	1500
AACAAGGGCA	TGGAACATCT	GCTCAACATG	AAGTGCAAAA	ATGTGCTCC	AGTGTATGAC	1560
CTGCTGCTGG	AGATGCTGAA	TGCCCAGCTG	CTTGGGGGT	GCAAGTCTC	CATCAAGGGG	1620
TCCGAGTGCA	GCCCGGCAGA	GGACAGTAAA	AGCAAGAGG	GCTCCAGSAA	CCCACACTCT	1680
CAGTGA						1686

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 530 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Asp	Ile	Lys	Asn	Ser	Pro	Ser	Ser	Leu	Asn	Ser	Pro	Ser	Ser	Tyr
1				5					10					15	
Asn	Gly	Ser	Gln	Ser	Ile	Leu	Pro	Leu	Gln	His	Gly	Ser	Ile	Tyr	Ile
		20						25					30		
Pro	Ser	Ser	Tyr	Val	Asp	Ser	His	His	Gln	Tyr	Pro	Ala	Met	Thr	Phe
		35					40					45			
Tyr	Ser	Pro	Ala	Val	Met	Asn	Tyr	Ser	Ile	Pro	Ser	Asn	Val	Thr	Asn
						55									
Leu	Gln	Gly	Gly	Pro	Gly	Arg	Gln	Thr	Thr	Ser	Pro	Asn	Val	Leu	Tyr
65				70				75						80	
Pro	Thr	Pro	Gly	His	Leu	Ser	Pro	Leu	Val	Val	His	Arg	Gln	Leu	Ser
			85					90						95	
His	Leu	Tyr	Ala	Gln	Pro	Gln	Lys	Ser	Pro	Tyr	Tyr	Gln	Ala	Arg	Ser
						105							110		
Leu	Gln	His	Thr	Leu	Pro	Val	Asn	Arg	Gln	Thr	Leu	Lys	Arg	Lys	Val
		115				120						125			
Ser	Gly	Asn	Arg	Tyr	Ala	Ser	Pro	Val	Thr	Gly	Pro	Gly	Ser	Lys	Arg

13	135	140
Asp Ala His Phe Cys	Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His	
145	150	155
Pyr Gly Val Trp Ser Cys	Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser	
165	170	175
Ile Gln Gly His Asn Asp Tyr Ile Cys	Pro Ala Thr Asn Gln Cys Thr	
180	185	190
Ile Asp Lys Asn Arg Arg Lys Ser Cys	Gln Ala Cys Arg Leu Arg Lys	
195	200	205
Cys Tyr Glu Val Gly Met Val Lys Cys	Gly Ser Arg Arg Glu Arg Cys	
210	215	220
Gly Tyr Arg Leu Val Arg Arg Gln Arg	Ser Ala Asp Glu Gln Leu His	
225	230	235
Cys Ala Gly Lys Ala Lys Arg Ser Gly	Gly His Ala Pro Arg Val Arg	
245	250	255
Glu Leu Leu Leu Asp Ala Leu Ser Pro	Glu Gln Leu Val Leu Thr Leu	
260	265	270
Leu Glu Ala Glu Pro Pro His Val Leu	Ile Ser Arg Pro Ser Ala Pro	
275	280	285
Phe Thr Glu Ala Ser Met Met Met Ser	Leu Thr Lys Leu Ala Asp Lys	
290	295	300
Glu Leu Val His Met Ile Ser Trp Ala	Lys Lys Ile Pro Gly Phe Val	
305	310	315
Glu Leu Ser Leu Phe Asp Gln Val Arg	Leu Leu Glu Ser Cys Trp Met	
325	330	335
Glu Val Leu Met Met Gly Leu Met Trp	Arg Ser Ile Asp His Pro Gly	
340	345	350
Lys Leu Ile Phe Ala Pro Asp Leu Val	Ile Asp Arg Asp Glu Gly Lys	
355	360	365
Cys Val Glu Gly Ile Leu Glu Ile Phe	Asp Met Leu Leu Ala Thr Thr	
370	375	380
Ser Arg Phe Arg Glu Leu Lys Leu Gln	His Lys Glu Tyr Leu Cys Val	
385	390	395
Lys Ala Met Ile Leu Leu Asn Ser Ser	Met Tyr Pro Leu Val Thr Ala	
405	410	415
Thr Gln Asp Ala Asp Ser Ser Arg Lys	Leu Ala His Leu Leu Asn Ala	
420	425	430
Val Thr Asp Ala Leu Val Trp Val Ile	Ala Lys Ser Gly Ile Ser Ser	
435	440	445
Gln Gln Gln Ser Met Arg Leu Ala Asn	Leu Leu Met Leu Leu Ser His	
450	455	460
Val Arg His Ala Ser Asn Lys Gly Met	Glu His Leu Leu Asn Met Lys	
465	470	475
Cys Lys Asn Val Val Pro Val Tyr Asp	Leu Leu Leu Glu Met Leu Asn	
485	490	495
Ala His Val Leu Arg Gly Cys Lys Ser	Ser Ile Thr Gly Ser Glu Cys	
500	505	510
Ser Pro Ala Glu Asp Ser Lys Ser Lys	Gln Gly Ser Gln Asn Pro Gln	
515	520	525
Ser His		
530		

1. INFORMATION FOR SEQ ID NO:3:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Thr His Tyr Ser Pro Ala Val Met Asn Tyr Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gly Tyr Ser Pro Ala Val Met Asn Tyr Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGAGCTTGT CGACCATCAT GACCGCTAT AGCCTGCTG TGATG

45

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGATCTAGAG TCGACCTGTC ACTGAGATG AGGCTCTGG

40

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Thr His Tyr

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (B) CLONE: R61010-2.24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CCATCACCGG CTAT

14

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: pOR2 sequencing primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTGGTAAGTT TAGTCTTTT GTC

23

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: oligo #12908

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCTTCACACG AAGGACTCTT TTGAG

25

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: oligo #13671

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GTTGATATA AATATTTT GTTTT

26

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Adaptor primer 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCATCCTAAT ACBACTCACT ATAGGGC

27

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 0 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Nested adaptor primer 2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ACTCACTATAGGGCTCGAGCGGC